

# Computing the Expected Hitting Time for the $n$ -Urn Ehrenfest Model via Two Methods

Sai SONG\* and Qiang YAO†

November 1, 2016

## Abstract

We study an  $n$ -urn version of the Ehrenfest model, where  $n \geq 2$ . In this model, there are  $n$  urns which are denoted from Urn 1 to Urn  $n$ . At the beginning,  $M$  balls are randomly placed in the  $n$  urns according to some law. Then at each time, one ball is chosen at random, removed from the current urn it resides in, and placed in one of the other  $n - 1$  urns equally likely. We use two methods to compute the expected hitting time when all  $M$  balls are in Urn 2 given that initially all  $M$  balls are in Urn 1.

**2010 MR subject classification:** 60C05, 60J10

**Key words:** Ehrenfest urn model, Markov chain, random walk, hitting time

## 1 Introduction

We extend the classical two-urn Ehrenfest model to the  $n$ -urn case, where  $n \geq 2$ . In this model, there are  $n$  urns which are denoted from Urn 1 to Urn  $n$ . At the beginning,  $M$  balls are randomly placed in the  $n$  urns according to some law. Then at each time, one ball is chosen at random, removed from the current urn it resides in, and placed in one of the other  $n - 1$  urns with the same probability. This model can be treated as a symmetric simple random walk on the graph  $G = (V, E)$ , where  $V = \{1, \dots, n\}^M$ , and  $E$  contains edges connecting two vertices in  $V$  if only one of their components differs. Therefore,  $G$  is a regular graph with  $n^M$  vertices, and each vertex has common degree  $(n - 1)M$ . Strictly speaking, if we let  $X_t = (X_t^{(1)}, \dots, X_t^{(M)})$  be the state at

---

\*School of Statistics, East China Normal University, Shanghai 200241, China.

†Corresponding author. School of Statistics, East China Normal University, Shanghai 200241, China. E-mail: qyao@sfs.ecnu.edu.cn.

time  $t = 0, 1, \dots$ , where  $X_t^{(i)}$  is the number of the urn in which the  $i$ th ball resides at time  $t$ , then  $\{X_t : t = 0, 1, \dots\}$  is a time homogeneous Markov chain on  $V$  with transition probability

$$p_{(x_1, \dots, x_M), (y_1, \dots, y_M)} = \begin{cases} \frac{1}{(n-1)M} & \text{if there exists } i \text{ s.t. } x_i \neq y_i, \text{ and } x_j = y_j \text{ for } j \neq i, \\ 0 & \text{otherwise.} \end{cases} \quad (1.1)$$

For  $x_1, \dots, x_M \in \{1, 2, \dots, n\}$ , denote by

$$T_{(x_1, \dots, x_M)} = \inf\{t \geq 0 : X_t = (x_1, \dots, x_M)\}$$

the first time that  $\{X_t\}$  hits state  $(x_1, \dots, x_M)$ . We are interested in

$$s_M = E(T_{(2, 2, \dots, 2)} \mid X_0 = (1, 1, \dots, 1)), \quad (1.2)$$

that is, the expected hitting time to the state that all  $M$  balls are in Urn 2 given that initially all  $M$  balls are in Urn 1. Note that we use the subscript “ $M$ ” to emphasize the number of balls because we will let it change in our second method. We have the following result.

**Theorem 1.1**  $s_M = \frac{(n-1)M}{n} \sum_{k=1}^M \frac{n^k}{k}.$

In Sections 2 and 3, we will give two totally different methods to prove Theorem 1.1, and in Section 4 we will give some concluding remarks.

## 2 Method I: Using an auxiliary Markov chain

In this section, we prove Theorem 1.1 with the aid of an auxiliary Markov chain. The idea is enlightened by Blom [1] and Lathrop et al [3].

### 2.1 Introduction to the auxiliary Markov chain

For any  $t = 0, 1, \dots$ , denote by  $S_t$  the number of balls in Urn 2 at time  $t$ . It is easy to see that  $\{S_t : t = 0, 1, \dots\}$  is a Markov chain on  $\{0, 1, \dots, M\}$  with transition probability

$$\begin{cases} p_{k, k-1} = \frac{k}{M} & k = 1, 2, \dots, M, \\ p_{kk} = \frac{(n-2)(M-k)}{(n-1)M} & k = 0, 1, \dots, M, \\ p_{k, k+1} = \frac{M-k}{(n-1)M} & k = 0, 1, \dots, M-1. \end{cases} \quad (2.1)$$

For  $k = 0, 1, \dots, M$ , define

$$\tau_k := \inf\{t \geq 0 : S_t = k\}.$$

It is easy to see that if  $S_0 = 0$ , then  $0 = \tau_0 < \tau_1 < \dots < \tau_M$ . Next, for  $k = 0, 1, \dots, M-1$ , define

$$e_k = E_0(\tau_{k+1} - \tau_k) := E(\tau_{k+1} - \tau_k \mid S_0 = 0). \quad (2.2)$$

The following proposition illustrates an important connection between the original process  $\{X_t\}$  and the auxiliary process  $\{S_t\}$ .

**Proposition 2.1**  $s_M = E_0(\tau_M) = \sum_{k=0}^{M-1} e_k$ .

Intuitively, Proposition 2.1 is resulted from the symmetric property. To verify this symmetric property strictly, we need the following lemma.

**Lemma 2.1** *For any  $a_1, \dots, a_M \neq 2$ , there exists an automorphism  $\phi$  on  $V$  such that*

$$\phi((2, 2, \dots, 2)) = (2, 2, \dots, 2), \quad \phi((1, 1, \dots, 1)) = (a_1, a_2, \dots, a_M).$$

**Proof.** For  $i = 1, \dots, M$ , define mapping  $\varphi_i : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$  in the way that

$$\varphi_i(c_i) = \begin{cases} a_i & \text{if } c_i = 1, \\ 1 & \text{if } c_i = a_i, \\ c_i & \text{otherwise.} \end{cases}$$

Then for any  $i = 1, \dots, M$ ,  $\varphi_i$  is a bijection on  $\{1, \dots, n\}$  with  $\varphi_i^{-1} = \varphi_i$ . Next, define

$$\begin{aligned} \phi : \quad V &\longrightarrow V \\ (c_1, \dots, c_M) &\longmapsto (\varphi_1(c_1), \dots, \varphi_M(c_M)) \end{aligned}$$

Then  $\phi$  is a bijection on  $V$  with  $\phi^{-1} = \phi$ , and

$$\phi((2, 2, \dots, 2)) = (2, 2, \dots, 2), \quad \phi((1, 1, \dots, 1)) = (a_1, a_2, \dots, a_M).$$

Furthermore, since  $\varphi_i$  is a bijection on  $\{1, \dots, n\}$  for  $i = 1, \dots, M$ , we have

$$\begin{aligned} (c_1, \dots, c_M) &\sim (d_1, \dots, d_M) \\ \iff \exists i \in \{1, \dots, M\} \text{ s.t. } c_i \neq d_i, \text{ and } c_j = d_j \text{ for } j \neq i \\ \iff \exists i \in \{1, \dots, M\} \text{ s.t. } \varphi_i(c_i) \neq \varphi_i(d_i), \text{ and } \varphi_j(c_j) = \varphi_j(d_j) \text{ for } j \neq i \\ \iff \phi((c_1, \dots, c_M)) &\sim \phi((d_1, \dots, d_M)) \end{aligned}$$

for any  $(c_1, \dots, c_M), (d_1, \dots, d_M) \in V$ . Therefore,  $\phi$  is an automorphism on  $V$  satisfying all what we need, as desired.  $\square$

**Proof of Proposition 2.1.** Obviously, by definition, we have

$$\tau_M = T_{(2,2,\dots,2)} \quad (2.3)$$

and

$$\{S_0 = 0\} = \sum_{a_1, \dots, a_M \neq 2} \{X_0 = (a_1, \dots, a_M)\}. \quad (2.4)$$

Since the transition probability for the random walk  $\{X_t\}$  defined by (1.1) satisfies

$$P_{(x_1, \dots, x_M), (y_1, \dots, y_M)} = P_{(y_1, \dots, y_M), (x_1, \dots, x_M)}$$

for any  $(x_1, \dots, x_M), (y_1, \dots, y_M) \in V$ , we can see that  $\{X_t\}$  is a reversible Markov chain. Therefore, for any  $a_1, a_2, \dots, a_M \neq 2$  and  $t \geq M$ ,

$$P(T_{(2,2,\dots,2)} = t \mid X_0 = (a_1, a_2, \dots, a_M)) = P(T_{(a_1, a_2, \dots, a_M)} = t \mid X_0 = (2, 2, \dots, 2)). \quad (2.5)$$

Especially,

$$P(T_{(2,2,\dots,2)} = t \mid X_0 = (1, 1, \dots, 1)) = P(T_{(1,1,\dots,1)} = t \mid X_0 = (2, 2, \dots, 2)). \quad (2.6)$$

From Lemma 2.1 combined with the fact that  $\{X_t\}$  is a symmetric random walk, we know that for any  $a_1, a_2, \dots, a_M \neq 2$ ,

$$P(T_{(a_1, a_2, \dots, a_M)} = t \mid X_0 = (2, 2, \dots, 2)) = P(T_{(1,1,\dots,1)} = t \mid X_0 = (2, 2, \dots, 2)). \quad (2.7)$$

(2.5), (2.6) and (2.7) lead to

$$P(T_{(2,2,\dots,2)} = t \mid X_0 = (1, 1, \dots, 1)) = P(T_{(2,2,\dots,2)} = t \mid X_0 = (a_1, a_2, \dots, a_M)) \quad (2.8)$$

for any  $a_1, \dots, a_M \neq 2$ . Denote by  $p$  the common value of both terms in (2.8), then  $p > 0$ , and

$$P(T_{(2,2,\dots,2)} = t, X_0 = (a_1, a_2, \dots, a_M)) = p \cdot P(X_0 = (a_1, a_2, \dots, a_M)) \quad (2.9)$$

for any  $a_1, \dots, a_M \neq 2$ . Therefore, by (2.3), (2.4) and (2.9), we obtain

$$\begin{aligned} P(T_{(2,2,\dots,2)} = t, S_0 = 0) &= \sum_{a_1, a_2, \dots, a_M \neq 2} P(T_{(2,2,\dots,2)} = t, X_0 = (a_1, a_2, \dots, a_M)) \\ &= p \cdot \sum_{a_1, a_2, \dots, a_M \neq 2} P(X_0 = (a_1, a_2, \dots, a_M)) = p \cdot P(S_0 = 0), \end{aligned}$$

and

$$P(\tau_M = t \mid S_0 = 0) = p = P(T_{(2,2,\dots,2)} = t \mid X_0 = (1, 1, \dots, 1))$$

for any  $t \geq M$ . Consequently,

$$\begin{aligned} s_M &= E(T_{(2,2,\dots,2)} \mid X_0 = (1, 1, \dots, 1)) = \sum_{t=M}^{\infty} t \cdot P(T_{(2,2,\dots,2)} = t \mid X_0 = (1, 1, \dots, 1)) \\ &= \sum_{t=M}^{\infty} t \cdot P(\tau_M = t \mid S_0 = 0) = E(\tau_M \mid S_0 = 0) = E_0(\tau_M). \end{aligned}$$

Furthermore, by (2.2) we have  $E_0(\tau_M) = \sum_{k=0}^{M-1} e_k$ . This completes the proof of Proposition 2.1.  $\square$

## 2.2 Computation of $e_k$

**Proposition 2.2** *For any  $k = 0, 1, \dots, M-1$ , we have  $e_k = \frac{(n-1)^{k+1}}{C_{M-1}^k} \sum_{l=0}^k \frac{C_M^l}{(n-1)^l}$ .*

**Proof.** If  $S_0 = 0$ , then  $\tau_1$  follows the geometric distribution with parameter  $\frac{1}{n-1}$ . So  $e_0 = E_0(\tau_1) = n-1$ . If  $k \geq 1$ , by the transition probability (2.1), we have

$$\begin{aligned} e_k &= \frac{k}{M} \cdot (e_{k-1} + e_k + 1) + \frac{(n-2)(M-k)}{(n-1)M} \cdot (e_k + 1) + \frac{M-k}{(n-1)M} \\ &= 1 + \frac{k}{M} \cdot e_{k-1} + \frac{(n-2)M+k}{(n-1)M} \cdot e_k. \end{aligned}$$

Therefore,  $\{e_k : k = 0, \dots, M-1\}$  follows the induction equation

$$\begin{cases} e_k = \frac{(n-1)k}{M-k} \cdot e_{k-1} + \frac{(n-1)M}{M-k} & k = 1, 2, \dots, M-1, \\ e_0 = n-1. \end{cases} \quad (2.10)$$

(2.10) is equivalent to

$$\begin{aligned} &\frac{e_k}{(n-1)^k k! (M-k-1)!} \\ &= \frac{e_{k-1}}{(n-1)^{k-1} (k-1)! (M-(k-1)-1)!} + \frac{M}{(n-1)^{k-1} k! (M-k)!} \quad k = 1, 2, \dots, M-1. \end{aligned}$$

So if we let  $f_{-1} = 0$  and  $f_l = \frac{e_l}{(n-1)^l l! (M-l-1)!}$  for  $l = 0, 1, \dots, M-1$ , then

$$f_l - f_{l-1} = \frac{M}{(n-1)^{l-1} l! (M-l)!}$$

for  $l = 0, 1, \dots, M-1$ . Therefore, for  $k = 0, 1, \dots, M-1$ , we have

$$f_k = \sum_{l=0}^k (f_l - f_{l-1}) = M \sum_{l=0}^k \frac{1}{(n-1)^{l-1} l! (M-l)!},$$

and

$$e_k = M(n-1)^k k! (M-k-1)! \sum_{l=0}^k \frac{1}{(n-1)^{l-1} l! (M-l)!} = \frac{(n-1)^{k+1}}{C_{M-1}^k} \sum_{l=0}^k \frac{C_M^l}{(n-1)^l},$$

as desired.  $\square$

### 2.3 Proof of Theorem 1.1 via Method I

By Proposition 2.2 together with the definition of beta and gamma functions, we have

$$\begin{aligned}
e_k &= \frac{(n-1)^{k+1}}{C_{M-1}^k} \sum_{l=0}^k \frac{C_M^{k-l}}{(n-1)^{k-l}} = \frac{n-1}{C_{M-1}^k} \sum_{l=0}^k (n-1)^l C_M^{k-l} \\
&= (n-1) \sum_{l=0}^k (n-1)^l \cdot \frac{Mk!(M-k-1)!}{(k-l)!(M-k+l)!} \\
&= (n-1)M \sum_{l=0}^k (n-1)^l \cdot C_k^l \cdot \frac{\Gamma(l+1)\Gamma(M-k)}{\Gamma(M-k+l+1)} \\
&= (n-1)M \sum_{l=0}^k (n-1)^l \cdot C_k^l \cdot B(M-k, l+1) \\
&= (n-1)M \sum_{l=0}^k (n-1)^l \cdot C_k^l \cdot \int_0^1 x^{M-k-1} (1-x)^l dx \\
&= (n-1)M \cdot \int_0^1 x^{M-k-1} [n - (n-1)x]^k dx
\end{aligned}$$

for  $k = 0, 1, \dots, M-1$ . The last equality is due to the identity

$$\sum_{l=0}^k C_k^l \cdot [(n-1)(1-x)]^l = [1 + (n-1)(1-x)]^k = [n - (n-1)x]^k.$$

Next, since

$$\sum_{k=0}^{M-1} x^{M-k-1} [n - (n-1)x]^k = \frac{x^M - [n - (n-1)x]^M}{n(x-1)},$$

we get

$$\begin{aligned}
\sum_{k=0}^{M-1} e_k &= (n-1)M \cdot \int_0^1 \frac{x^M - [n - (n-1)x]^M}{n(x-1)} dx \\
&\stackrel{x=1-t}{=} \frac{(n-1)M}{n} \cdot \int_0^1 \frac{[1 + (n-1)t]^M - (1-t)^M}{t} dt. \tag{2.11}
\end{aligned}$$

For  $k = 0, 1, \dots, M$ , define

$$g_k = \int_0^1 \frac{[1 + (n-1)t]^k - (1-t)^k}{t} dt,$$

then we have  $g_0 = 0$  and

$$\begin{aligned}
g_k &= \int_0^1 \frac{[1 + (n-1)t]^{k-1} - (1-t)^{k-1} + (n-1)t[1 + (n-1)t]^{k-1} + t(1-t)^{k-1}}{t} dt \\
&= g_{k-1} + \int_0^1 (n-1)[1 + (n-1)t]^{k-1} dt + \int_0^1 (1-t)^{k-1} dt
\end{aligned}$$

$$\begin{aligned}
& \frac{u=1+(n-1)t}{s=1-t} g_{k-1} + \int_1^n u^{k-1} du + \int_0^1 s^{k-1} ds \\
& = g_{k-1} + \int_0^n u^{k-1} du = g_{k-1} + \frac{n^k}{k}
\end{aligned}$$

for  $k = 1, 2, \dots, M$ . Therefore, by (2.11) we get

$$\sum_{k=0}^{M-1} e_k = \frac{(n-1)M}{n} \cdot g_M = \frac{(n-1)M}{n} \sum_{k=1}^M (g_k - g_{k-1}) = \frac{(n-1)M}{n} \sum_{k=1}^M \frac{n^k}{k}.$$

Together with Proposition 2.1, we get the desired result.  $\square$

### 3 Method II: Using stopping times

In this section, we utilize some stopping times to prove Theorem 1.1. Now we fix the number of urns ( $n$ ) and let the number of balls change, we will use the notation  $s_k$  instead of  $s_M$  defined in (1.2) if we consider  $k$  balls. Our main step is to prove the following proposition, which gives an induction formula for  $\{s_k : k = 1, 2, \dots\}$ .

**Proposition 3.1**  $\{s_k : k = 1, 2, \dots\}$  satisfies the induction formula

$$\begin{cases} s_k = \frac{k}{k-1} \cdot s_{k-1} + (n-1)n^{k-1} & k = 2, 3, \dots, \\ s_1 = n-1. \end{cases} \quad (3.1)$$

#### 3.1 Proof of Proposition 3.1

When  $k \geq 2$ , if the number of balls is  $k$ , then the state space of  $\{X_t\}$  becomes  $\{1, \dots, n\}^k$ , and the transition probability is

$$\begin{aligned}
& P_{(x_1, \dots, x_k), (y_1, \dots, y_k)} \\
& = \begin{cases} \frac{1}{(n-1)k} & \text{if there exists } i \text{ s.t. } x_i \neq y_i, \text{ and } x_j = y_j \text{ for } j \neq i, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

For any  $x_1, \dots, x_{k-1} \in \{1, \dots, n\}$ , define

$$A_{(x_1, \dots, x_{k-1})} := \{(x_1, \dots, x_{k-1}, x_k) : x_k \in \{1, \dots, n\}\}.$$

That is,  $A_{(x_1, \dots, x_{k-1})}$  contains all points in  $V$  such that the first  $k-1$  components are  $x_1, \dots, x_{k-1}$ .

We next define several stopping times. First, denote by

$$T = \inf\{t > 0 : X_t = (2, 2, \dots, 2)\}.$$

Then

$$s_k = E_{(1,1,\dots,1)}(T). \quad (3.2)$$

Next, we define a sequence of stopping times  $\{\tau_k : k = 0, 1, 2, \dots\}$  inductively by  $\tau_0 = 0$  and

$$\tau_k = \inf\{t > \tau_{k-1} : X_t \in A_{(2,2,\dots,2)}\}$$

for  $k \geq 1$ . Clearly, we have

$$P_{(1,1,\dots,1)}\left(\bigcup_{k=1}^{\infty}\{T = \tau_k\}\right) = \sum_{k=1}^{\infty} P_{(1,1,\dots,1)}(T = \tau_k) = 1. \quad (3.3)$$

We first introduce several lemmas before proving Proposition 3.1.

**Lemma 3.1** *We have*

$$E_{(1,1,\dots,1)}(\tau_i - \tau_{i-1}) = \begin{cases} \frac{k}{k-1} \cdot s_{k-1} & \text{if } i = 1, \\ n^{k-1} & \text{if } i \geq 2. \end{cases}$$

**Proof.** We define two auxiliary Markov chains on  $\{1, \dots, n\}^{k-1}$ . First, denote

$$Y_t := (X_t^{(1)}, \dots, X_t^{(k-1)})$$

for  $t = 0, 1, 2, \dots$ . Then  $\{Y_t\}$  is a Markov chain which illustrates the positions of the first  $k-1$  balls with transition probability

$$p_{xy} = \begin{cases} \frac{1}{k} & \text{if } y = x, \\ \frac{1}{k(n-1)} & \text{if } y \sim x, \\ 0 & \text{otherwise.} \end{cases}$$

Here the notation “ $y \sim x$ ” means that  $y$  is a neighbor of  $x$ .

Next, denote by  $\{Z_t\}$  the Markov chain on  $\{1, \dots, n\}^{k-1}$  which illustrates the process with  $k-1$  balls and  $n$  urns with transition probability

$$q_{xy} = \begin{cases} \frac{1}{(k-1)(n-1)} & \text{if } y \sim x, \\ 0 & \text{otherwise.} \end{cases}$$

For  $x \in \{1, \dots, n\}^{k-1}$ , let  $f_x = E_x(\tau_1)$  for  $\{Y_t\}$  (that is, under the transition probability  $\{p_{xy}\}$ ). Then let  $g_x = E_x(\tau_1)$  for  $\{Z_t\}$  (that is, under the transition probability  $\{q_{xy}\}$ ). By the transition probability  $\{q_{xy}\}$ , we can get that  $\{g_x\}$  satisfies

$$g_x = \begin{cases} 1 + \frac{1}{(k-1)(n-1)} \sum_{y \sim x} g_y & \text{if } x \neq (2, \dots, 2), \\ 0 & \text{if } x = (2, \dots, 2). \end{cases} \quad (3.4)$$



Similarly, by the transition probability  $\{p_{xy}\}$ , we can get that  $\{f_x\}$  satisfies

$$f_x = \begin{cases} 1 + \frac{1}{k} \cdot f_x + \frac{1}{k(n-1)} \sum_{y \sim x} f_y & \text{if } x \neq (2, \dots, 2), \\ 0 & \text{if } x = (2, \dots, 2). \end{cases} \quad (3.5)$$

Note that (3.5) can be written as

$$\frac{k-1}{k} f_x = \begin{cases} 1 + \frac{1}{(k-1)(n-1)} \sum_{y \sim x} \frac{k-1}{k} f_y & \text{if } x \neq (2, \dots, 2), \\ 0 & \text{if } x = (2, \dots, 2). \end{cases} \quad (3.6)$$

Comparing (3.6) with (3.4), we can see that  $\{g_x\}$  and  $\left\{\frac{k-1}{k} f_x\right\}$  obey the same difference equation and have the same initial value. Therefore,  $g_x = \frac{k-1}{k} f_x$  for any  $x \in \{1, \dots, n\}^{k-1}$ . Especially, we have  $f_{(1, \dots, 1)} = \frac{k}{k-1} g_{(1, \dots, 1)}$ .

Since  $E_{(1, \dots, 1, a)}(\tau_1)$  is the same for any  $a \in \{1, \dots, n\}$ , we have  $f_{(1, \dots, 1)} = E_{(1, 1, \dots, 1)}(\tau_1)$ . Together with the fact that  $g_{(1, \dots, 1)} = s_{k-1}$ , we get  $E_{(1, 1, \dots, 1)}(\tau_1) = \frac{k}{k-1} \cdot s_{k-1}$ .

Furthermore, since  $\{Y_t\}$  is a reversible Markov chain on  $\{1, \dots, n\}^{k-1}$  (with  $n^{k-1}$  vertices), there exists a unique invariant distribution which puts equally like mass  $\frac{1}{n^{k-1}}$  on the  $n^{k-1}$  vertices. So for any  $i = 2, 3, \dots$ , we have  $E_{(1, 1, \dots, 1)}(\tau_i - \tau_{i-1}) = n^{k-1}$ , as desired.  $\square$

**Lemma 3.2**  $P_{(1, 1, \dots, 1)}(T = \tau_1) = P_{(2, 2, \dots, 2)}(T = \tau_1)$ .

**Proof.** For  $i = 1, 2, \dots, k$ , denote

$$B_{2i-1} = \left\{ (x_1, \dots, x_k) \in V : \sum_{l=1}^{k-1} \mathbf{1}_{\{x_l=2\}} = i-1, x_k \neq 2 \right\},$$

$$B_{2i} = \left\{ (x_1, \dots, x_k) \in V : \sum_{l=1}^{k-1} \mathbf{1}_{\{x_l=2\}} = i-1, x_k = 2 \right\},$$

where  $\mathbf{1}_{\{\cdot\}}$  is the indicator function. Then  $V = \sum_{m=1}^{2k} B_m$ . By the symmetric property which is similar to Lemma 2.1,  $P_x(T = \tau_1)$  is the same for the  $x$ 's belonging to the same  $B_m$ . So for any  $m = 1, \dots, 2k$ , we have

$$P_x(T = \tau_1) = P_{B_m}(T = \tau_1)$$

for any  $x \in B_m$ . Denote  $p_m = P_{B_m}(T = \tau_1)$  for  $m = 1, \dots, 2k$ . Then

$$P_{(1, \dots, 1)}(T = \tau_1) = p_1, \quad P_{(2, \dots, 2)}(T = \tau_1) = p_{2k}.$$

We next prove  $p_1 = p_{2k}$ . For any  $m_1, m_2 \in \{1, \dots, 2k\}$ , let

$$q_{m_1 m_2} = P(X_1 \in B_{m_2} \mid X_0 \in B_{m_1}),$$

which equals to  $P_x(X_1 \in B_{m_2})$  for any  $x \in B_{m_1}$  by the symmetric property which is similar to Lemma 2.1. Then we have

$$\left\{ \begin{array}{l} q_{2i, 2i-1} = \frac{1}{k} \\ q_{2i, 2i-2} = \frac{k-i}{i-1} \\ q_{2i, 2i+2} = \frac{k-i}{k} \cdot \frac{1}{n-2} \\ q_{2i, 2i} = \frac{k-i}{k} \cdot \frac{n-1}{n-2} \\ q_{2i-1, 2i} = \frac{1}{k} \cdot \frac{n-1}{k-i} \\ q_{2i-1, 2i+1} = \frac{k-i}{k} \cdot \frac{1}{n-1} \\ q_{2i-1, 2i-3} = \frac{i-1}{k} \\ q_{2i-1, 2i-1} = \frac{k-i+1}{k} \cdot \frac{n-2}{n-1} \end{array} \right. \quad (1 \leq i \leq k),$$

and  $q_{m_1 m_2} = 0$  otherwise. From this, we first get

$$p_{2k} = q_{2k, 2k-2} p_{2k-2} = \frac{k-1}{k} p_{2k-2}. \quad (3.7)$$

Next,  $\{p_i : i = 1, \dots, 2k-2\}$  follows

$$\left\{ \begin{array}{l} p_1 = \frac{n-2}{n-1} p_1 + \frac{1}{k} \cdot \frac{1}{n-1} p_2 + \frac{k-1}{k} \cdot \frac{1}{n-1} p_3, \\ p_2 = \frac{1}{k} p_1 + \frac{k-1}{k} \cdot \frac{1}{n-2} p_2 + \frac{k-1}{k} \cdot \frac{1}{n-1} p_4, \\ p_3 = \frac{1}{k} p_1 + \frac{k-1}{k} \cdot \frac{1}{n-1} p_3 + \frac{1}{k} \cdot \frac{1}{n-1} p_4 + \frac{k-2}{k} \cdot \frac{1}{n-1} p_5, \\ \dots, \\ p_{2k-4} = \frac{k-3}{k} p_{2k-6} + \frac{1}{k} p_{2k-5} + \frac{2}{k} \cdot \frac{n-2}{n-1} p_{2k-4} + \frac{2}{k} \cdot \frac{1}{n-1} p_{2k-2}, \\ p_{2k-3} = \frac{k-2}{k} p_{2k-5} + \frac{2}{k} \cdot \frac{n-2}{n-1} p_{2k-3} + \frac{2}{k} \cdot \frac{1}{n-1} p_{2k-2}. \end{array} \right. \quad (3.8)$$

Note that we do not need  $p_{2k-1}$  in the last inequality since if  $\{X_t\}$  touches  $B_{2k-1}$  before  $B_{2k}$ , then  $T > \tau_1$ . The first equation in (3.8) implies

$$p_1 = \frac{1}{k} p_2 + \frac{k-1}{k} p_3. \quad (3.9)$$

The second and third equations in (3.8) imply

$$\left(1 - \frac{k-1}{k} \cdot \frac{n-2}{n-1}\right) p_2 = \frac{1}{k} p_1 + \frac{k-1}{k} \cdot \frac{1}{n-1} p_4 \quad (3.10)$$

and

$$\left(1 - \frac{k-1}{k} \cdot \frac{n-2}{n-1}\right) p_3 = \frac{1}{k} p_1 + \frac{1}{k} \cdot \frac{1}{n-1} p_4 + \frac{k-2}{k} \cdot \frac{1}{n-1} p_5. \quad (3.11)$$

Putting (3.10) and (3.11) into (3.9), we get

$$\begin{aligned} & k \left(1 - \frac{k-1}{k} \cdot \frac{n-2}{n-1}\right) p_1 \\ &= \left(\frac{1}{k} p_1 + \frac{k-1}{k} \cdot \frac{1}{n-1} p_4\right) + (k-1) \left(\frac{1}{k} p_1 + \frac{1}{k} \cdot \frac{1}{n-1} p_4 + \frac{k-2}{k} \cdot \frac{1}{n-1} p_5\right). \end{aligned}$$

That is,

$$p_1 = \frac{2}{k} p_4 + \frac{k-2}{k} p_5.$$

Similarly, we can inductively get

$$p_1 = \frac{1}{k} p_2 + \frac{k-1}{k} p_3 = \frac{2}{k} p_4 + \frac{k-2}{k} p_5 = \dots = \frac{k-1}{k} p_{2k-2}. \quad (3.12)$$

The last equality is due to the absence of  $p_{2k-1}$  as explained above. From (3.7) and (3.12), we get  $p_1 = p_{2k}$ , as desired.  $\square$

**Remark 3.1** *In fact, we can deduce from (3.8) that  $P_{(1,1,\dots,1)}(T = \tau_1) = P_{(2,2,\dots,2)}(T = \tau_1) = \frac{n^{k-1} - 1}{n^k - 1}$ . The procedure is much more complicated, and is omitted here.*

**Lemma 3.3** *For any  $i = 1, 2, \dots$  and  $t \geq 0$ , we have*

$$P_{(1,1,\dots,1)}(T - \tau_{i+1} = t \mid T > \tau_i) = P_{(2,\dots,2,1)}(T - \tau_1 = t).$$

*That is,  $\{(T - \tau_{i+1} \mid T > \tau_i)\}$  have the same distribution for  $i = 1, 2, \dots$ .*

**Proof.** For any  $i = 1, 2, \dots$ , we have  $\{T > \tau_i\} = \bigcup_{x \neq 2} \{X_{\tau_i} = (2, \dots, 2, x)\}$ . From the strong Markov property, we have for any  $t \geq 0$ ,  $i = 1, 2, \dots$  and  $x \neq 2$ ,

$$\begin{aligned} & P_{(1,1,\dots,1)}(T - \tau_{i+1} = t \mid X_{\tau_i} = (2, \dots, 2, x)) \\ &= \sum_{m=1}^{\infty} P_{(1,1,\dots,1)}(T - \tau_i = m + t, \tau_{i+1} - \tau_i = m \mid X_{\tau_i} = (2, \dots, 2, x)) \\ &= \sum_{m=1}^{\infty} P_{(2,\dots,2,x)}(T = m + t, \tau_1 = m) \\ &= P_{(2,\dots,2,x)}(T - \tau_1 = t) = P_{(2,\dots,2,1)}(T - \tau_1 = t). \end{aligned}$$

The last equality is due to the symmetric property which is similar to Lemma 2.1. Therefore,

$$P_{(1,1,\dots,1)}(T - \tau_{i+1} = t \mid T > \tau_i) = P_{(2,\dots,2,1)}(T - \tau_1 = t)$$

for any  $i = 1, 2, \dots$  and  $t \geq 0$ , as desired.  $\square$

**Lemma 3.4** Let  $p = P_{(1,1,\dots,1)}(T > \tau_1)$  and  $q = P_{(1,1,\dots,1)}(T > \tau_2 \mid T > \tau_1)$ , then we have

$$\frac{p}{1-q} = n-1.$$

**Proof.** By the above lemmas, we have

$$\begin{aligned} p &= P_{(2,2,\dots,2)}(T > \tau_1) = \sum_{x \neq 2} P_{(2,2,\dots,2)}(X_{\tau_1} = (2, \dots, 2, x)) \\ &= (n-1)P_{(2,2,\dots,2)}(X_{\tau_1} = (2, \dots, 2, 1)) \\ &= (n-1)P_{(2,2,\dots,1)}(X_{\tau_1} = (2, \dots, 2, 2)) \\ &= (n-1)P_{(2,2,\dots,1)}(T = \tau_1) = (n-1)(1-q). \end{aligned}$$

The first equality comes from Lemma 3.2, the third and fourth equalities are due to the symmetric property which is similar to Lemma 2.1, and the last equality comes from Lemma 3.3, as desired.  $\square$

Now we are ready to prove Proposition 3.1.

**Proof of Proposition 3.1.** When  $k = 1$ ,  $T$  follows the geometric distribution with probability  $\frac{1}{n-1}$ . Therefore,  $s_1 = n-1$ .

When  $k \geq 2$ , let  $x = E_{(1,1,\dots,1)}(T - \tau_1)$ ,  $y = E_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_1)$ . Then by Lemma 3.3,  $E_{(1,1,\dots,1)}(T - \tau_3 \mid T > \tau_2) = y$ . Similar to the proof of Lemma 3.3, we have

$$E_{(1,1,\dots,1)}(T - \tau_1 \mid X_{\tau_1} = (2, \dots, 2, x)) = \begin{cases} E_{(1,1,\dots,1)}(T - \tau_1 \mid T > \tau_1) & \text{if } x \neq 2, \\ 0 & \text{if } x = 2. \end{cases}$$

Therefore,

$$\begin{aligned} E_{(1,1,\dots,1)}(T - \tau_1) &= \sum_{x \neq 2} E_{(1,1,\dots,1)}(T - \tau_1 \mid X_{\tau_1} = (2, \dots, 2, x)) \cdot P(X_{\tau_1} = (2, \dots, 2, x)) \\ &= pE_{(1,1,\dots,1)}(T - \tau_1 \mid T > \tau_1), \end{aligned} \tag{3.13}$$

where  $p = P_{(1,1,\dots,1)}(T > \tau_1)$  is defined as in Lemma 3.4. Similarly, we can prove

$$E_{(1,1,\dots,1)}(\tau_2 - \tau_1 \mid T > \tau_1) = E_{(1,1,\dots,1)}(\tau_2 - \tau_1) = n^{k-1} \tag{3.14}$$

and

$$E_{(1,1,\dots,1)}(\tau_3 - \tau_2 \mid T > \tau_2) = E_{(1,1,\dots,1)}(\tau_3 - \tau_2) = n^{k-1} \tag{3.15}$$

together with the result of Lemma 3.1. Also, we can prove

$$E_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_1) = qE_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_2), \tag{3.16}$$

where  $q = P_{(1,1,\dots,1)}(T > \tau_2 \mid T > \tau_1) = P_{(1,1,\dots,1)}(T > \tau_2)$  is defined as in Lemma 3.4. By (3.13) and (3.14), we have

$$\begin{aligned} x &= E_{(1,1,\dots,1)}(T - \tau_1) = pE_{(1,1,\dots,1)}(T - \tau_1 \mid T > \tau_1) \\ &= p[E_{(1,1,\dots,1)}(\tau_2 - \tau_1 \mid T > \tau_1) + E_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_1)] \\ &= p(n^{k-1} + y). \end{aligned} \tag{3.17}$$

Similarly, by (3.15) and (3.16), we have

$$\begin{aligned} y &= E_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_1) = qE_{(1,1,\dots,1)}(T - \tau_2 \mid T > \tau_2) \\ &= q[E_{(1,1,\dots,1)}(\tau_3 - \tau_2 \mid T > \tau_2) + E_{(1,1,\dots,1)}(T - \tau_3 \mid T > \tau_2)] \\ &= q(n^{k-1} + y). \end{aligned} \tag{3.18}$$

By (3.17) and (3.18), together with the result of Lemma 3.4, we get

$$x = \frac{p}{1-q} n^{k-1} = (n-1)n^{k-1}.$$

Then together with Lemma 3.1, we get

$$s_k = E_{(1,1,\dots,1)}(T) = E_{(1,1,\dots,1)}(\tau_1) + E_{(1,1,\dots,1)}(T - \tau_1) = \frac{k}{k-1} \cdot s_{k-1} + (n-1)n^{k-1}$$

for  $k \geq 2$ , as desired.  $\square$

### 3.2 Proof of Theorem 1.1 via Method II

If we let  $h_0 = 0$  and  $h_k = \frac{s_k}{k}$  for  $k = 1, 2, \dots$ , then by (3.1),

$$h_k - h_{k-1} = \frac{n-1}{n} \cdot \frac{n^k}{k}$$

for  $k = 1, 2, \dots$ . Therefore, we have

$$h_M = \sum_{k=1}^M (h_k - h_{k-1}) = \frac{n-1}{n} \sum_{k=1}^M \frac{n^k}{k},$$

and  $s_M = \frac{(n-1)M}{n} \sum_{k=1}^M \frac{n^k}{k}$ , as desired.  $\square$

## 4 Concluding remarks

In this paper, we use two methods to compute the expected hitting time when all  $M$  balls are in one urn given that initially all  $M$  balls are in another urn. The first method, which utilizes the auxiliary Markov chain, is easy to comprehend, since the auxiliary chain is 1-dimensional. However, it does not illustrate the induction relationship (3.1). The second method, which utilizes a series of stopping times, makes a better illustration for it.

There may be some other available methods. For example, by using electric networks. Readers can refer to Doyle and Snell [2] or Lyons and Peres [4] for the introduction of this method. Palacios [5] used the method of electric networks to consider the 2-urn case. However, it may become difficult to use the “Y-Delta” transformation to simplify the network when  $n$  is large, especially when  $n \geq 4$ . It is an interesting topic to be considered in future works.

Also, in this paper, we only consider the “unbiased” case, that is, the ball is chosen randomly and put in other urns randomly with equal probabilities. It will be interesting to consider the case of biased probabilities or preferential probabilities in future works.

**Acknowledgment.** The second author was partially supported by the Natural Science Foundation of China (No.11201150, No.11126236), the Natural Science Foundation of Shanghai (16ZR1409700), and the 111 Project (B14019).

## References

- [1] Blom, G. (1989). Mean transition times for the Ehrenfest urn model. *Adv. Appl. Prob.* **21**, 479–480.
- [2] Doyle, P.G. and Snell, E.J. (1984). Random walks and electric networks. *Carus Math. Monographs*, **22**, Math. Assoc. Amer., Washington, D.C.
- [3] Lathrop, E.D., Goldstein, I.H. and Chen, Y.P. (2016). A note on a generalized Ehrenfest urn model: another look at the mean transition times. *J. Appl. Prob.* **53**, 630–632.
- [4] Lyons, R. and Peres, Y. (2016). Probability on trees and networks. Cambridge University Press. To appear, available at <http://pages.iu.edu/~rdlyons/prbtree/prbtree.html>.
- [5] Palacios, J.L. (1994). Another look at the Ehrenfest urn via electric networks. *Adv. Appl. Prob.* **26**, 820–824.